

## PRODUCTION TRAINING EFFECTS ON /r/-l/ PERCEPTION SKILLS OF NATIVE ADULT JAPANESE SPEAKERS

Chris D. Martz

Indiana University

### Introduction

The problems experienced by non-native speakers in perceiving or producing non-native phonemic contrasts is well-documented (Flege, 1988; Goto, 1971). Sheldon and Strange (1982:244) provide a concise description of this problem as it pertains to Japanese learners of English and their difficulty in distinguishing between /r/ and /l/:

Since Japanese has one liquid phoneme and English has two, Japanese speakers must learn to perceive the distinction between English /r/ and /l/. This task is complicated by the fact that the variants of Japanese /r/ are phonetically unlike English /r/, although one variant appears to be phonetically similar to English /l/. Japanese learners of English have to learn to perceive and to produce an entirely new type of consonant, namely approximant /r/, and they have to learn to differentiate it from /l/. However, since the Japanese /r/-like variant is not identical to any of the allophones of English /l/, further perceptual and articulatory refinement is necessary.

Miyawaki et al. (1975) characterize the Japanese /r/ as having two allophones: 1) in the syllable-initial position, a loose alveolar stop; 2) in the intervocalic position, a flap. In addition, Dissosway-Huff (1981) points out that, whether word initial or intervocalic, the Japanese /r/ is strictly syllable-initial, which precludes its use in initial clusters, final clusters, or the word-final position. When a L2 phonemic contrast is not present in a learner's L1, it creates perception and production problems that may be difficult to overcome (Strange and Dittmann, 1984; Jamieson and Morosan, 1986). Indeed, one must overcome previous linguistic experience if a person is to perceive phonemic distinctions (Dissosway-Huff, 1981).

The difficulty of /r/-l/ discrimination is, however, open to debate. Goto (1971) and Miyawaki et al. (1975) contend that phonemic discrimination (both perception and production) must take place at an early age since Japanese adults in their studies displayed obvious discrimination difficulties. Pisoni et al. (1982), on the other hand, found that native English-speaking adults were able to quickly learn to distinguish three different types of stops (i.e., voiced, voiceless unaspirated, and voiceless aspirated), even though aspiration is not phonologically distinctive in English. They postulated that early training in one's L1 does not permanently change perceptual/sensory mechanisms of a listener; these mechanisms can be retrained to perceive new distinctions found in other languages. This seems to be indirectly supported by Dissosway-Huff's (1981) observation that one of the three Japanese subjects in her study seemed to dramatically improve /r/-l/ perceptual discrimination skill with only regular exposure to stimuli, without any feedback. In other words, mere exposure and practice may have been enough for this subject to "retrain" distinctive boundaries.

Dissosway-Huff (1981) states that there is a theoretical assumption often taken by second language acquisition scholars that perception precedes production. After all, it seems only

logical that one cannot produce a sound distinction without first being able to perceive it; furthermore, production is an "active skill", and thus takes more practice to master than perception, a "passive skill". Nevertheless, several studies seem to contradict this assumption in that performance competence is often shown to be superior to that of perception, specifically with reference to Japanese learners of English and their ability to distinguish between English /r/ and /l/. Goto (1971) found that those subjects whose English ability was best had much better production than perception skills of /r/ and /l/. According to Dissosway-Huff (1981), low-intermediate Japanese ESL learners also exhibited superior production skills as compared to their ability to perceptually discriminate English /r/ and /l/. Sheldon and Strange (1982) found that six Japanese subjects attending an American university had production capabilities that exceeded their perceptual skills. These subjects also found word-initial clusters to be the most difficult to accurately perceive, and word-final singletons the easiest. Finally, Yamada et al. (1994) showed that Japanese subjects who performed poorly on an /r, l, w/ identification task also had intelligibility of production scores that varied from low to high; however, those subjects with higher performances on the identification task also tended to score high on production intelligibility. This suggests that, at least for some subjects, production skills exceeded perception skills, but the reverse was never the case.

Much research has tried to determine the most effective language domain (i.e., perception vs. production) for the training of language contrasts. Catford and Pisoni (1970) explored whether exclusive articulatory training or auditory training of adults resulted in the most accurate production of exotic language sounds. They found that subjects who received articulatory training were able to produce previously unknown sounds accurately more than twice as often as subjects who had received auditory training. In addition, the articulatory-trained group identified sounds significantly more accurately than the auditory group. Both Goto (1971) and Sheldon and Strange (1982) seem to concur with the superiority of articulatory training to facilitate accurate production of unknown sounds. Goto observed that his Japanese subjects with the best English /r/ and /l/ production seemed to rely on "kinetic sensations" of their articulators to produce the sounds correctly, while Sheldon and Strange noted that their good producers of English /r/ and /l/ had been taught by means of articulatory training rather than auditory cues.

Recent studies, however, have explored the effects of perception training on the perception of non-native phonemic contrasts, and use of different speakers and phonemes in different word positions, creating stimulus variability. Logan et al. (1991) introduced stimulus variability by using /r/-/l/ perception training stimuli employing different word environments and spoken by several talkers. Using the training stimuli in a forced-choice training identification task with feedback, they found that subjects showed improvement in /r/-/l/ minimal-pair identification, and did so in less response time.

Two experiments were conducted in a study by Lively et al. (1993). The first focused on the effects of /r/-/l/ discrimination training in specific phonetic environments with multiple talkers. Adopting the conclusions of Atkinson (1972) as to the perceptual difficulty of specific phonetic environments, Lively et al. trained their subjects in the three most difficult environments: initial singleton, initial cluster, and intervocalic, omitting training in word-final positions. Results indicated that subjects improved in /r/-/l/ identification and response time for all word positions from pretest to posttest. This improvement generalized to new words, a new

talker, and to untrained (word-final) positions. A second experiment had subjects similarly trained, but with words spoken by just one talker. Although subjects improved during training in pretest/posttest performance, they did not generalize to tokens spoken by a new talker. Both experiments indicate the value of perceptual training that includes stimulus variability in the form of a variety of talkers producing tokens that use target phonemes in different word environments. Such stimuli can help to develop robust categorization skills for new phonemic contrasts not found in a learner's L1.

Magnuson et al. (1995) found similar results with reference to talker variability. Five groups of subjects were trained with five different talkers, each group trained with tokens from only one talker. Only two of the five groups showed generalization to new tokens and talkers. The three other groups failed to show any perceptual learning as indicated by pretest/posttest results. Thus, single-talker training brought about less than satisfactory perceptual results.

Of particular interest to the present study are: 1) the length of training needed to achieve improved performance, and 2) "cross-domain" effects. Bradlow et al. (1995) explored the effects of training in one domain (perception) on performance in the same (perception) and different (production) domains. Eleven monolingual adult native speakers of Japanese were pretested in both perception and production domains. The perception pretest consisted of a list of /r/-/l/ minimal-pair words originally used by Strange and Dittmann (1984), with /r/ and /l/ in four phonetic environments: initial singleton, initial cluster, intervocalic, and final singleton. The production pretest consisted of a separate list /r/-/l/ minimal-pair words covering the four phonetic environments of the perception pretest, plus initial triple clusters, medial clusters, and final clusters. The training phase consisted of 45 sessions over 3-4 weeks of perceptual identification training with feedback. Training stimuli were /r/-/l/ minimal-pair words covering five phonetic environments (initial singleton, initial cluster, intervocalic, final singleton, final cluster), and spoken by five native speakers of English. The posttest phase included a perceptual identification posttest (same words as pretest) and two perception generalization tests: one with new words spoken by a new talker, and the other with new words (different words than those of the first generalization test) spoken by one of the talkers who had recorded the training stimuli (a "familiar" speaker). The production posttest was the same as the production pretest. A control group of 12 subjects took all tests concurrently with the experimental group, but did not receive perceptual training. Results showed that subjects who had received the perception training performed significantly better than the control group on both perception and production tasks.

Similar research using the production domain for training is needed. The purpose of this study was to test the following questions: 1) Can similar "cross-domain" discrimination effects be seen when subjects are trained in the production domain? In other words, can production training positively affect perceptual skills? 2) Can minimal production training still show positive discrimination effects? Results may increase our understanding of the most effective training domain for phonemic contrasts, as well as the most efficient training duration. Although production data were also gathered in this study, the focus of this paper will be the "cross-domain" effects of /r/-/l/ production training on perception skills.

## Methodology

### *Subjects*

The participants in this study were native Japanese undergraduate students of Tsuru University, Tsuru, Japan, who were involved in a special one-month English program. They were attending classes at the Intensive English Program/Center for English Language Training, Indiana University, Bloomington, Indiana. These students were chosen because of their limited exposure to spoken English, given their length of stay for the aforementioned program. All of the participants (three males, three females) had had no more than two weeks stay in an English-speaking country, and were 20-21 years of age. All had begun learning English in middle school, had had little or no listening/speaking training, and all exhibited varying degrees of /r/-/l/ discrimination problems. Hearing was reported to be within normal limits.

### *Testing Instruments/Procedures*

A pretest and posttest were administered before and after training. The pretest and posttest were derived from a list of minimal pairs originally created by Strange and Dittmann (1984), and subsequently used by Bradlow et al. (1995). One minimal pair was deleted from the list because the words contained both /r/ and /l/, which made their use unreliable given the testing method employed that eliminated the use of orthography and required a subject to choose only one of the phonemes. The words were randomly arranged into two lists by a random-listing Microsoft Excel '97 program. The first randomized list became the pretest, and the second randomized list the posttest. Two generalization tests were also administered after training. Generalization (GEN) tests #1 (new words, new speaker) and #2 (new words, old speaker) were taken directly from the Bradlow et al. (1995) study. Once again, due to the specific testing protocol, minimal pair words containing both /r/ and /l/ were eliminated from the generalization lists. Randomization of the two word lists was achieved using the aforementioned Excel program.

Recording of the test lists and perception testing were accomplished using a Panasonic SV-3800 DAT recorder, Symetrix pre-amplifier, and Sony MDR-V6 digital headphones. The pretest and posttest lists, the GEN #1 list, and the GEN #2 list were recorded by three native English speakers, all of whom were from the midwestern United States, and had no distinct dialectal or regional accent. The pretest and posttest lists were recorded by one speaker, the GEN #1 list was recorded by a different speaker not previously encountered by the subjects (new speaker), and the GEN #2 list was recorded by a speaker familiar to the subjects, the training instructor (old speaker). To insure reliability of the recordings, all four recorded lists were tested by having them administered to two native speakers of English. Both native speakers agreed with all intended test items, as well as with each other, except for one word (*tile*) from the pretest, which was subsequently dropped from the pretest list.

Testing procedures differed from those of Bradlow et al., whose subjects were tested with a two-alternative, forced-choice minimal word pair identification procedure first developed by Jamieson and Morosan (1986) and later used by Logan et al. (1991). In the Bradlow et al. study, a subject was shown an orthographic representation of a minimal word pair on a computer

screen; the spoken test word was then presented to the subject through headphones, and the subject would indicate the word on the screen that he/she had heard. In contrast, to prevent any possible selection bias due the use of orthography, subjects of the present study received scoring sheets for which the choices “r”, “l”, and “neither” were written for each test item. They were instructed to circle the appropriate response corresponding to the sound heard for each presented word. Both pretest and posttest were made up of words containing /r/, /l/, or neither phoneme; generalization tests had either /r/ or /l/ words. Although the “neither” response was a viable choice for only the pretest and posttest, it was presented for all tests to lessen the likelihood of correct guessing between only two alternatives. The entire study consisted of pretesting (1 session), training (3 sessions), and posttesting/generalization testing (2 sessions) conducted over a 10-day period. Posttesting/generalization testing was administered over two sessions to control for any fatigue factor.

### *Training Procedure*

The training list was comprised of words used for training by Bradlow et al. However, whereas the Bradlow study trained subjects with words containing /r/-/l/ in all word positions, the present study chose to limit word positions for training to initial singleton and initial cluster. This decision was made to check for positional training effects on perception/production. Once again, the words were randomized in the same manner as the test items to mix phoneme type and word position. Appendix D shows the randomized training list used.

Training consisted of three 45-minute sessions over three days. Subjects were first instructed with the help of modeling on the alveolar placement of the tongue for the /l/ sound, and drilled with words from the training list. Following this, subjects were instructed with the help of modeling on the tongue placement for the velar /r/ sound. This particular /r/ choice was made to maximize tongue placement difference between the /l/ and /r/ phonemes. The subjects were taken from a known sound, /i/, then shown how to glide the sides of the tongue farther back along the upper teeth to produce /r/. This method concurs with the idea of Catford and Pisoni (1970) that effective articulation training is composed of small steps from known articulatory postures and movements to new/unknown postures and movements. Words from the training list were used for practicing /r/. All subjects demonstrated the ability to produce both phonemes accurately in isolation and in words by the end of the first session. Subsequent sessions were used to review placement principles, and practice production skills using words from the training list.

### *Results*

Due to an unequal number of words per test, phoneme type, and phoneme word position, percentage correct was used as the dependent variable. Alpha level for all analyses was set at .05. A one-way ANOVA was performed comparing the total percentages correct of the four perception tests (pretest, posttest, GEN #1, GEN #2). Although correct percentages ranged from 65.11 (Pretest) to 73.85 (GEN #1), no significant differences were found among the four tests ( $F_{(3,15)} = 1.251$ ;  $p = .326$ ).

A second ANOVA was performed on pretest/posttest results to determine if position of the /r-/ phonemes within the words (initial singleton [IS], initial cluster [IC], intervocalic [VCV], final singleton [FS]) and/or the particular phoneme type (/r/ or /l/) might have affected perception performance. Only phoneme type showed significant results ( $F_{(1,75)} = 6.44$ ;  $p = .0133$ ), indicating that, across both tests, participants were more likely to perceive /r/ correctly than /l/. Although word position was not significant ( $F_{(3,75)} = 1.34$ ;  $p = .2688$ ), the trained positions (IS and IC) both showed improvement from pretest to posttest for both /r/ and /l/.

A final ANOVA was conducted to determine if position and/or phoneme type affected perception on the two generalization tests. Position main effect was significant ( $F_{(4,90)} = 3.059$ ;  $p = .021$ ), indicating that a final phoneme word positions (FS and FC) were more correctly identified than initial word positions (IS and IC). Position\*test interaction was also significant ( $F_{(4,90)} = 2.580$ ;  $p = .043$ ), reflecting the effect of so few VCV items on correct/error percentages. Few VCV tokens would result in large percentage shifts caused by only a single correct or incorrect response. Position\*phoneme type interaction was highly significant ( $F_{(4,90)} = 6.711$ ;  $p < .001$ ), indicating that specific phonemes in specific word positions were more correctly identified. In this case, /r/ was more often correctly identified than /l/ in the IS, FS, and FC positions, while /l/ fared better than /r/ in the IC and VCV positions. Test\*phoneme type interaction was also significant ( $F_{(1,90)} = 5.819$ ;  $p = .018$ ), indicating differences in correct answer percentages for the two phonemes between the two tests; this was especially evident for GEN #1: (GEN #1: /r/ = 80.53%, /l/ = 67.17%; GEN #2: /r/ = 71.19%, /l/ = 69.62%).

Finally, of interest were word-final (FS and FC) perception errors that occurred on the generalization tests. Of 138 total FS/FC items across all subjects in GEN #1, 37 errors were committed, 10 of which were answers of "neither" instead of correct responses of /r/ or /l/. In other words, for 27% of the errors made on words in which /r/ or /l/ were in these final positions, the subjects perceived neither phoneme. The results are even more striking in GEN #2 results: of 132 total FS/FC items across all subjects, 31 errors were committed, 22 of which were answers of "neither" instead of /r/ or /l/. In this instance, subject inability to perceive a present /r/ or /l/ constituted 71% of the errors found when these phonemes were in word-final positions. In contrast, combined pretest and posttest results showed only 20 errors out of 90 word-final items, only 1 of which was an error of the "neither" type (5% of the errors). Individual differences among subjects show that the number of "neither" errors attributed to GEN #2 is due largely to one subject (TK), who accounted for 10 of the 31 "neither" errors. The remaining errors were more evenly distributed among the other five subjects. Of interest also is the fact that all 10 GEN #1 errors were committed when /l/ was presented in the final position, but more than twice as many GEN #2 errors were committed when /r/ (15) was presented than /l/ (7).

To summarize the findings, minimal production training did not significantly affect overall perception of /r/ and /l/ across the four tests. Pretest/posttest analysis revealed that phoneme type was a significant factor, with /r/ identified correctly more often than /l/ across the two tests. In addition, identification of /r-/l/ improved in the trained positions (IS and IC) from pretest to posttest, although this difference was not statistically significant.

The generalization tests indicated that phoneme word position was a significant factor, indicating that final word positions (FS and FC) were more correctly identified than initial positions (IS and IC). A significant position\*test interaction reflected the paucity of VCV

tokens, and the consequent percentage-correct shifts resulting from minimal error changes. A highly significant position\*phoneme type interaction resulted from /r/ being more correctly identified in the IS, FS, and FC positions, whereas /l/ was more correctly identified in the IC and VCV positions. A significant test\* phoneme type interaction resulted from the differences in percentages correct for /r/ and /l/ (especially for GEN #1), with /r/ correctly identified more often than /l/ across both tests. Finally, a "neither phenomenon" was observed for word-final positions. Several times for word-final positions (FS and FC) of both tests, subjects did not hear a presented /r/ or /l/ sound. Phonemes missed seemed to differ with the different speakers used for the two tests. This "neither phenomenon" did not occur in the pretest or posttest.

## Discussion

The tenuous nature of any findings from this study must be readily acknowledged due to the limited number of subjects. It must be remembered, however, these subjects were chosen specifically because of their lack of native English exposure, a factor that may well have been replicated only in Japan. The most important finding of this study may be the realization that several variables need to be considered in future research. In addition, one should consider the methodological and statistical shortcomings that occurred in this study as a result of adopting testing/statistical procedures used by a previous study.

Pretest and posttest results indicated that phoneme type was indeed a factor affecting correct responses, /r/ being perceived correctly more often than /l/. Generalization test results indicated that phoneme position within a word was also a factor affecting correct responses. An unexpected finding concerning phoneme word position was that percentages correct for non-trained positions (VCV = 81.48%; FS = 81.91%; FC = 68.08%) were generally higher than trained positions (IS = 70.67%; IC = 61.07%). This can partially be explained by the fewer VCV, FS, and FC items used in comparison to the IS and IC items on these two test lists; however, the reason for discrepancies between percentages correct for word-initial and word-final singletons and clusters is less clear. This finding, however, is consistent with the assertion of Lively et al. (1993) and Atkinson (1972) that initial positions are among the most difficult phonetic environments (along with VCV) for Japanese subjects to contrast /r/ and /l/. Hardison (1997) also found that native-Japanese speakers had more difficulty perceiving /r/-/l/ correctly in IS and IC positions than FS and FC positions. The fact that the position\*test interaction was significant ( $p = .043$ ) can also be attributed to the low number of VCV items. So few items can result in a great percentage correct change brought about by even a single error.

Generally, results showed that differences in perception may be affected by what phoneme is presented (/r/ or /l/), and in what word position. Such conclusions, however, are suspect because care was not taken to have equal numbers of words containing each phoneme in all word positions. Closest to meeting these criteria were the pretest and posttest, but even with these lists, numbers of presented words were not equal per position, phoneme type, or total items per test.

The generalization tests were the greatest violators of this position/phoneme equality principle. Granted that the original intent of the GEN #1 and GEN #2 tests in the Bradlow et al.

(1995) study was to merely determine the effect of speaker (familiar vs. unfamiliar) on generalizability to new words; however, not controlling for the number of words per phoneme in each word position makes any claims about speaker effect highly suspect. In other words, if phoneme type and/or phonemes per word position are not comparably represented among the test items, any resulting differences could just as easily be attributable to these uncontrolled variables as to the speaker variable. In addition, the two tests are not comparable because the test items themselves and phoneme word-position counts were different for both tests.

To complicate variable control further, recent research indicates that vowel environment may also affect /r/-/l/ phoneme recognition. Hardison (1997) found that /e, ε/ were the easiest vowel environments in which to perceive /r/ and /l/, and /o/ the most difficult. Vowel\*position interaction was also highly significant ( $p = .0001$ ), indicating the need for variable control and the complexity of variable interrelationships.

Finally, Bradlow et al. used the generalization tests on both treatment and control groups, allowing a means of comparison for these "post-treatment" tests. However, even assuming the validity of this comparison, one cannot assume that the treatment group's post-training performance actually indicates generalization in the absence of similar testing prior to treatment. Concerning the present study, the absence of a control group probably did little to further weaken methodology already fraught with problems.

Amid these confounding variables is one that was not considered by either this or the Bradlow study: the effect of a subject's knowing the meaning/spelling of only one word of a minimal pair on the answer choice. In other words, is a subject more likely to "hear" a word with which they are familiar than its unknown minimal-pair counterpart? Word familiarity (semantic or orthographic) may indeed exert an unintentional, yet powerful influence upon what is and is not perceived by non-native subjects.

The "neither" phenomenon alluded to previously--subjects perceiving neither /r/ nor /l/ in the FS and FC positions--was indeed a surprising finding. Of course, the fact that Japanese /r/ is always syllable-initial and only found word-initially and intervocalically (Dissosway-Huff, 1981; Miyawaki et al., 1975) may have contributed to this performance anomaly. In addition, perhaps the minimal training received by the subjects had started the shifting of categorical boundaries of perception for /r/ and /l/; consequently, such /r/-/l/ boundaries in a state of flux caused perceptual confusion when subjects were confronted with new words containing these phonemes in the untrained FS/FC positions. It is likely that perceptual category shifts require more time to stabilize than was allowed for this study--indeed, the Bradlow et al. study used 45 training sessions over a period of 3-4 weeks before finding significant perception changes. An additional factor may have been that native-English production of /r/ and /l/ is not uniform across word positions (Lehiste, 1964; Hagiwara, 1995). It could be that the word-final productions were different enough from what the subjects had learned in the word-initial positions of their training that it caused them to not perceive the sounds altogether. Perhaps these position-sensitive production differences were greatest for the GEN #2 (familiar) speaker. During training, the subjects may have "templated" the acoustical /r/-/l/ properties of the familiar speaker/trainer, but only for the trained word-initial positions. Subsequent exposure to the same individual producing /r/-/l/ in different word positions than those heard during training may have



created perceptual "confusion", causing the subjects to not perceive the phonemes in the untrained word-final positions.

Spectrographic analysis of the taped speakers might shed light on the contribution of acoustical initial vs. final /r/-/l/ differences to this curious lack of word-final perception. This may indeed be the case since the types of errors (neither for /r/ vs. neither for /l/) differed greatly for the two speakers of the tests. The work of Lehiste (1964) and Hagiwara (1995) indicate significant differences between /r/ and /l/, including such factors as: 1) word-initial vs word-final position; 2) formant dependence on/independence from the adjacent vowel; 3) formant transition speed. In addition, formant cues used for distinguishing /r/ and /l/ may differ between native-Japanese and native-English speakers, with the former focusing on F2 and the latter on F3 (Hardison, 1997).

Future research should concern itself with: 1) creating methodologies that control as many extraneous variables as possible; 2) the effects of semantic/orthographic awareness on perception of minimal pairs; 3) what acoustical cues are salient for L1 (native) and L2 (nonnative) listeners in identifying contrastive L1 sounds; 4) establishing an optimal training domain and duration for teaching contrastive phonemes to L2 learners. Perhaps this study's greatest contribution is a caveat to future researchers: one must be mindful of the multitude of variables that may affect subject perceptions and test results. The following have been mentioned in this paper and are, by no means, presented as an exhaustive list—the effects of orthography, semantic/orthographic awareness, speaker, word position (and allophonic variations per word position), vowel environment, number of test items per contrastive phoneme/phoneme word position, perceptive difficulty per contrastive phoneme. Only when these and, likely, additional variables are incorporated into future research designs will a more accurate picture of L2 perception/production be realized.

## REFERENCES

- Atkinson, R. C. 1972. Ingredients for a theory of instruction. *American Psychologist* 27. 921-31.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Y. Tohkura. 1995. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. Indiana University, Bloomington. Distributed as Research on Spoken Language Processing: Progress Report No.20.
- Catford, J. C., and D. B. Pisoni. 1970. Auditory vs. articulatory training in exotic sounds. *The Modern Language Journal* 104. 477-81.
- Dissowsay-Huff, P. 1981. Japanese difficulty with English /l/ and /r/: Importance of both perception and production research in second language acquisition. Research in Phonetics, Report #2, Department of Linguistics, Indiana University, Bloomington, Indiana.
- Flege, J. E. 1988. The production and perception of foreign language speech sounds. In *Human communication and its disorders: A review*, ed. by H. Winitz, 224-401. Norwood, NJ: Ablex.

Goto, H. 1971. Auditory perceptions by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia* 9. 317-23.

Hagiwara, R. E. 1995. Acoustic realizations of American /r/ as produced by women and men. Distributed as UCLA Working Papers in Phonetics No. 90. Los Angeles, CA: UCLA dissertation.

Hardison, D. M. 1997. Bimodal input in second-language speech: Focus on /r/ and /l/. *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second-Language Speech*, ed. by J. Leather and A. James, 125-34. Klagenfurt, Austria: University of Klagenfurt.

Jamieson, D. G., and D. E. Morosan. 1986. Training non-native speech contrasts in adults: Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception and Psychophysics* 40. 205-15.

Lehiste, I. 1964. Acoustic characteristics of selected English consonants. *International Journal of American Linguistics* 30. 10-115.

Lively, S. E., Logan, J. D., and D. B. Pisoni. 1993. Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America* 94. 1242-55.

Logan, J. D., Lively, S. E., and D. B. Pisoni. 1991. Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America* 89. 874-86.

Magnuson, J. S., Yamada, R. A., Tohkura, Y., Pisoni, D. B., Lively, S. E., and A. R. Bradlow. 1995. The role of talker variability in non-native phoneme training. *Proceedings of the 1995 spring meeting of the Acoustical Society of Japan*. 393-94.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J., and O. Fujimura. 1975. An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics* 18. 331-40.

Pisoni, D. B., Aslin, R. N., Perey, A. J., and B. L. Hennessy. 1982. Some effects of laboratory training on identification and discrimination of voicing contrasts in stop consonants. *Journal of Experimental Psychology: Human Perception and Performance* 8. 297-314.

Sheldon, A., and W. Strange. 1982. The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3. 243-61.

Strange, W., and S. Dittmann. 1984. Effects of discrimination training on the perception of /r/-/l/ by Japanese adults learning English. *Perception and Psychophysics* 36. 131-45.

Yamada, R.A., Strange, W., Magnuson, J.S., Pruitt, J.S., and W. D. Clarke, III. 1994. The intelligibility of Japanese speakers' productions of American English /r/, /l/, and /w/, as evaluated by native speakers of American English. Proceedings of the International Conference of Spoken Language Processing, Yokohama, Japan.